January 2005

Measurement Equivalence of the Center for Epidemiological Studies Depression Scale for Latino and Anglo Adolescents: A National Study

Lisa J. Crockett
University of Nebraska-Lincoln, ecrockett1@unl.edu

Brandy A. Randall
North Dakota State University, Brandy.Randall@ndsu.edu

Yuh-Ling Shen
University of Nebraska-Lincoln

S T. Russell
strussell@arizona.edu

Anne K. Driscoll
University of California, Davis, akdriscoll@ucdavis.edu

Measurement Equivalence of the Center for Epidemiological Studies Depression Scale for Latino and Anglo Adolescents: A National Study

Lisa J. Crockett, *University of Nebraska–Lincoln*
Brandy A. Randall, *North Dakota State University*
Yuh-Ling Shen, *University of Nebraska–Lincoln*
Stephen T. Russell & Anne K. Driscoll, *University of California, Davis*

The cross-ethnic measurement equivalence of the Center for Epidemiologic Studies Depression Scale (CES-D; L. S. Radloff, 1977) was examined using a subsample of adolescents ($N = 10,691$) from the National Longitudinal Study of Adolescent Health, Configural and metric invariance, as well as functional and scalar equivalence, were examined for Anglo American, Mexican American, Cuban American, and Puerto Rican American youths age 12-18 years. Confirmatory factor analysis (CFA) in each group provided evidence of configural invariance for European and Mexican American adolescents but not for Cuban and Puerto Rican youths. A 2-group CFA for Anglo and Mexican Americans demonstrated partial metric invariance for these groups. Multigroup structural equation modeling indicated similar relations between CES-D scores and self-esteem for all 4 groups, supporting cross-ethnic functional and scalar equivalence. The results have implications for using the CES-D in cross-ethnic research and, more broadly, for the assessment and treatment of depression in Latinos.

Depression is one of the most common psychological problems experienced by adolescents. An estimated 5% of 9- to 17-year-olds suffer from major depression (Shaffer et al., 1996), and estimated rates among adolescents are as high as 8% (Birmaher et al., 1996; Garrison, Waller, Cuffe, & McKeown, 1997). An even greater number of youths experience symptoms of depression, especially depressed affect, but do not meet the criteria for a diagnosis of major depression (Shaffer et al., 1996).

In the past decade, increased interest has focused on ethnic differences in depression. Research has pointed to ethnic disparities in depression among adults (e.g., Minsky, Vega, Miskimen, Oara, & Escobar, 2003; Vega & Rumbaut, 1991) and more recently to ethnic differences among adolescents. Most studies show higher rates of depressive symptoms and disorders in Latino adolescents compared with Anglo adolescents (Knight; Virdin, Ocampo, & Roosa, 1994; Roberts & Chen, 1995; Roberts & Sobhan, 1992; Weinberg & Emslie, 1987), although a few show the opposite pattern (Vega, Khoury, Zimmerman, Oil, & Warheit, 1995). For example, Roberts, Roberts, and Chen (1997) found

that 12% of Latino adolescents met criteria for depression, compared with 6.3% of Anglo youths. In another recent study, Latinos reported higher symptom scores than Anglo Americans, African Americans, or Asian Americans (Seigel, Aneshensel, Taub, Cantwell, & Driscoll, 1998). Because Latinos now comprise the largest minority ethnic group in the United States, these differences warrant particular attention.

Comparative studies raise questions about the cross-ethnic validity of the depression measures used. Most mental health assessment instruments were initially developed and tested on samples comprised largely of European Americans; as a consequence, we do not know how well these measures assess depression in other ethnic-racial groups, including Latinos (Vega & Rumbaut, 1991). Equivalence of measures is crucial because if measures have differential meaning or differential validity for Latinos as compared with European Americans, prevalence estimates for Latinos would be inaccurate (Vandenberg & Lance, 2000) and group comparisons misleading (Hui & Triandis, 1985). However, to our knowledge, no published study has systematically examined the

equivalence of depression measures across subgroups of Latino adolescents. The present study sought to address this gap by using data from the National Longitudinal Study of Adolescent Health (Add Health), which included a version of the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff. 1977).

The Problem of Nonequivalent Measures

Lack of measurement equivalence can occur on several levels. For example. the construct of depression may differ across cultural groups. with each group conceptualizing it differently and using different symptoms to identify it. If so. a measure of depression developed for one group (e.g., Anglos) would fail to capture relevant aspects of the concept as understood by the other group (e.g.. Latinos). Even if the construct is the same. the measure used could assess that construct more poorly in one group than the other. If the items used to measure depression are poorer indicators of depression in one group. estimates for that group will be less accurate. Finally. for a given measure. different groups may interpret the possible responses differently or use the response scale differently. In such cases. a particular score could reflect different amounts of the construct in the two groups; for example. a score of 16 might represent severe depression in one group but only moderate depression in another.

The problem of nonequivalence of measures also extends to subgroups within a cultural population. Latinos are a heterogeneous group. representing distinct nationalities. cultural traditions. and levels of acculturation (U .S. Department of Health and Human Services. 2001). The cultural differences among subgroups (e.g.. Mexican, Cuban, and Puerto Rican Americans) could result in the differential expression of depression as well as group differences in prevalence rates (e.g., Choi, 2002). In fact. data from a large community survey of Latinos (The Hispanic Health and Nutrition Examination Survey) indicate subgroup differences in rates of depression. with Cuban Americans showing lower levels of depressive symptomatology than other Hispanic subgroups (Narrow. Rae. Moscicki. Locke. & Regier. 1990). However, in many studies of adolescent mental health. these groups are pooled.

The consequences of using nonequivalent measures are potentially serious. If a measure used to screen for cases of depression is valid and accurate for one group (e.g.. Anglos) but less so for another (e.g.. Latinos). then applying the standard cutoffs will lead to misclassification in the second group. resulting in false positives. false negatives, or both. On a national level. the use of differentially valid assessment instruments will distort prevalence estimates for depression in some groups arid yield inaccurate information on ethnic disparities in mental health. potentially leading to misguided policy initiatives. Regarding clinical practice, if presenting symptoms differ or some symptoms are more salient for one group than another. clinicians need to be attuned to these differences to provide accurate assessment and effective treatment (Choi, 2002; Sue, 1998).

Several converging lines of evidence point to the possibility that the experience and conceptualization of depression may differ across cultural groups. Cross-cultural research has raised the possibility that mental health and illness are contextually based and culturally embedded (e.g., Kleinman, 1986). According to the sociosomatic formulation in medical anthropology, "a person'

s context. ..influences the severity and type of symptoms experienced" (James & Prilleltensky, 2002, p. 1134); in addition, cultural categories may influence which symptoms are culturally acceptable. Thus, even though most cultures have concepts of sadness (a basic human emotion) and grief (a common human experience), they may not have a concept of depression as a mental illness. Even if they do, the symptoms associated with the core notion of sadness may differ. Furthermore, some syndromes may be linked to specific cultures (identified as culture-bound syndromes in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* [American Psychiatric Association, 1994]). To provide a concrete example, Mexican culture includes a concept of *nervios,* an emotional affliction that is related to anxiety and depression but is recognized as distinct (Salgado de Snyder, Diaz-Perez; & Ojeda, 2000). There is also considerable evidence that Latinos tend to somaticize mental health problems, reporting more physical symptoms of distress than European Americans (U.S. Department of Health and Human Services, 2001). Factor analytic studies have reinforced these clinical observations, showing that depressed affect and somatic symptoms load on the same factor for Latinos but on separate factors for Anglos (e.g., Roberts, 1980, 1992). Thus, particular symptoms of depression may be more closely linked in some cultures than in others.

Despite the possibility of measurement nonequivalence, relatively little multiethnic research has addressed this issue for Latinos. Rather, standard measures of adolescent depression, developed largely with European American samples, have been applied to Latinos (or subgroups of Latinos) without close attention to the potential impact of differential validity. Although a few studies have sought to examine the equivalence of depression measures for Latino and Anglo youths (e.g., Knight et al., 1994; Knight, Virdin, & Roosa, 1992), almost none have included representative samples of youths (for an exception, see Roberts, 1992). To our knowledge, the three largest subgroups of Latino youths in the United States (Mexican, Cuban, and Puerto Rican Americans) have not been systematically compared in a single study. The present study sought to address this gap by examining the measurement equivalence of a common measure of depression, the CES-D, for Anglo adolescents and three subgroups of Latino .youths: Mexican Americans, Cuban Americans, and Puerto Rican Americans.

Racial-Ethnic Differences in the CES-D

The CES-D was originally designed to measure somatic and affective symptoms of depression in community samples of adults (Radloff, 1977). Exploratory factor analysis (EFA) of a large community sample of Black and White adults yielded four factors in each group. The factors corresponded to Negative Affect, Positive Affect, Somatic Symptoms, and Interpersonal Symptoms. Since that initial study, the CES-D has been widely used for epidemiological and clinical studies (Orme, Reis, & Hen, 1986) with an increasingly diverse range of populations (e.g., Fountoulakis et al., 2001; Liang, Tran, Krause, & Markides, 1989), including adolescents (e.g., McArdle, Johnson, Hishinuma, Miyamoto, & Andrade, 2001).

Most cross-ethnic analyses of the CES-D factor structure have focused on adults; these studies provide mixed support for the original four-factor structure. Roberts, Vernon, and Rhoades (1989) found that a four-factor model fit for both Mexican Amer-

ican and European American psychiatric patients. However, other studies identified different sets of factors in Latino groups. In a pooled sample drawn from three studies of urban Latinos, Posner , Stewart, Marin, and Perez-Stable (2001) found that the four-factor confirmatory model showed a marginally acceptable fit; however, additional analyses indicated that this model fit for Latina women but not Latino men. In a study of Mexican Americans, the CES-D items "lonely," "sad," and "crying" loaded together on the Negative Affect scale, unlike results: with Anglo samples (Garcia & Marks, 1989). Similarly, in a comparison of African Americans, Anglo Americans, and Mexican Americans using EFA, some negative affect and somatic items loaded together for the Mexican Americans but not the other groups (Roberts, 1980). Finally, Golding and Aneshensel (1989) reported high conceptual equivalence of items among Anglo Americans and both U.S.-born and Mexico-born Mexican Americans but small differences in factor structure: The sleep disturbance item loaded on the Somatic factor for U.S.-born Mexican Americans, on the Negative Affect factor for Mexico-born Mexican Americans, and on both factors for Anglo Americans. These studies indicate that the four-factor solution fits in some cases but not others and fits better for African Americans and Anglo Americans than for Latino Americans. There appear to be subtle but interesting differences in factor structure between Anglo and Mexican Americans; less is known about other Latino groups.

Subgroup differences among Latinos have been documented as well. Guamaccia, Angel, and Worobey (1989) used confirmatory factor analysis (CFA) to examine the fit of the four-factor structure initially reported by Radloff (1977) for Mexican American, Cuban American, and Puerto Rican American adults. The authors concluded that the model did not fit the data in any of the three groups, but their conclusion was based solely on a significant chi-square test, which is easy to obtain with large samples (Kline, 1998). EFA for each of the three Latino subgroups supported a three-factor structure in each subgroup. However, there were subgroup differences in the factors on which specific items loaded. For instance, Cuban Americans showed a factor that was interpreted as reflecting feelings of rejection by and isolation from the larger society. These results underscore the importance of examining the CES-D within subgroups rather than treating Latinos as a single homogeneous group.

Studies of the CES-D in Latino Adolescents

The factor structure originally identified for Black and White adults has been replicated with non-Hispanic White adolescents (Roberts, Andrews, Lewinsohn, & Hops, 1990). However, replication with adolescents from other ethnic-racial groups has rarely been attempted, and few studies have examined the factor structure of the CES-D among Latino adolescents. Using EFA with a 12-item version of the CES-D, Roberts (1992) found three factors for Anglo, African American, Mexican American, and other Hispanic adolescents, corresponding to Positive Affect, Negative Affect, and Somatic symptoms. Although the Positive Affect factor was the same across groups, for Mexican and other Hispanic adolescents, there was a tendency for a few somatic items to load on the Negative Affect factor as well as the Somatic factor. Moreover, the factor intercorrelations varied across the four groups. For Anglos and African Americans, correlations between the Pos-

itive Affect factor and the other two factors were negative, but for Mexican and other Hispanic youths, all three factors were positively intercorrelated, suggesting different relations among dimensions of depression for Latino adolescents as compared with their Anglo and African American peers.

Measurement equivalence is typically assessed through the use of statistical analyses comparing the properties of a measure in two or more groups. Discussions in the cross-cultural literature consider several types of measurement equivalence (Hui & Triandis, 1985). The most basic type of equivalence is *conceptual equivalence*—unless two groups hold the same concept of depression, there is little purpose in determining whether measures of that construct are equally valid across groups (Hui & Triandis, 1985). Conceptual equivalence is supported by *configural invariance,* which is demonstrated when the items comprising a measure show the same factor structure in two groups (Ghorpade, Hattrup, & Lackritz, 1999; Vandenberg & Lance, 2000). It is also supported by *functional equivalence,* which is demonstrated when two behaviors are expressed in the same situations; serve the same purposes; and have similar antecedents, correlates, and consequents across cultures (Hui & Triandis, 1985). *Item equivalence* indicates that the items used to assess the construct are identical and have the same meaning for members of two cultural groups. It is supported by tests of *metric invariance,* which occurs when the factor loadings of items in the measure are invariant across the two groups (Vandenberg & Lance, 2000). In contrast, different factor loadings indicate that some items are better (more central) indicators of depression in one group than the other. Finally, *scalar equivalence* exists when "the construct is measured on the same metric" for both groups, such that a given value on the scale refers to the same "degree, intensity, or magnitude of the construct" in both groups (Hui & Triandis, 1985, p. 135). Functional and scalar equivalence are examined using regression techniques or structural equation models. Functional equivalence is supported by similarity of regression slopes across groups; scalar equivalence requires similarity of regression slopes and intercepts (Knight & Hill, 1998). It is important to note that configural and metric invariance focus on properties of the items comprising the measure, whereas functional and scalar equivalence, as examined here, focus on .associations between the scale as a whole and other theoretically related variables.

To date, studies of measurement equivalence with Latino adolescents have been restricted by sample limitations or by use of an abbreviated version of the CES-D. In addition, EFA has been used to examine factor structure rather than confirmatory techniques. Thus, the question of whether the original four-factor solution applies to Latino youths is unresolved. Furthermore, the issue of subgroup differences has rarely been addressed in studies of adolescents, even though data from studies of adults suggest that subgroup differences are likely. Perhaps most important, the functional and scalar equivalence of the CES-D across subgroups of Latino adolescents or between Latinos and Anglos have not been systematically examined. Such information is crucial if researchers are to appropriately use and interpret scores based on this measure.

The Present Study

The goal of the present study was to examine the equivalence of the CES-D for Anglo adolescents and three subgroups

of Latino youths. We examined configural and metric invariance, as well as functional and scalar equivalence, to address two core questions: (a) Do Anglo and Latino adolescents (Mexican, Cuban, and Puerto Rican American) appear to have the same concept of depression as measured by the CES-D? and (b) If so, does the CES-D measure depression equally well in all four groups, such that scores have the same meaning? Tests of configural invariance and functional equivalence were used to assess equivalence of constructs; tests of metric invariance and scalar equivalence addressed the comparability of scores.

To assess functional and scalar equivalence, we examined associations between scores on the CES-D and a theoretically related variable-self-esteem. Self-esteem was expected to be inversely associated with depression. A negative association between self-esteem and depressive symptoms has been reported in several studies (e.g., Aunola, Stattin, & Nunni, 2000; Ohannessian, Lerner, Lerner, & von Eye, 1999); however, the strength of this association could vary across ethnic groups.

Apart from the basic need for information on the equivalence of depression measures for Latinos, the importance of these analyses is twofold. First, Add Health is the most current nationally representative sample of adolescents; thus, estimates based on this data set are likely to be more accurate than results based on older cohorts or local community samples. Second, the Add Health study has become an important data source for understanding adolescent depression (e.g., Rushton, Forcier, & Schectman, 2002; van Dulman et al., 2002); evidence of measurement equivalence is critical for the accurate interpretation of resulting prevalence estimates for the full population as well as for ethnic subgroups.

Method

Sample

Add Health is a nationally representative study of U.S. adolescents in Grades 7 through 12 (Udry, 1998). The study was designed to examine the health status of adolescents, as well as influences on their health-related behaviors, with a focus on the multiple contexts in which adolescents live. Add Health used a multistage, stratified, school-based, cluster sampling design. Specifically, a sample of 80 high schools and 52 middle schools from the United States (132 schools) was selected with unequal probability of selection. The sample was selected to be representative of U.S. schools with respect to region of country, urbanicity, school type, ethnicity, and school size.

All students who completed an in-school questionnaire or who were listed on a school roster of one of the participating schools were eligible for the home interview. A representative sample of these youths (the core sample) was selected and supplemented with several special subsamples. Some ethnic groups were sampled in proportion to their size within the U.S. population; others (Chinese, Cuban, Puerto Rican) were oversampled to ensure adequate sample sizes for analysis (Bearman, Jones, & Udry, 1997). For the in-home survey, questionnaires were administered via laptop computer. Audio Computer Assisted Interviewing was used for sensitive questions. The analytic sample was drawn from the 20,745 adolescents who completed the first in-home survey (contractual data set). Adolescents who were between the ages of 12 and 18 years at Wave 1 of the in-home survey and had valid sample weights were identified. (If more than one adolescent in a family participated, one sibling was randomly selected for inclusion to eliminate nonindependent cases.) The present analyses included youths who self-identified as White, non-His-

panic (Anglo) or as Mexican, Cuban, or Puerto Rican Americans of any race. The final analytic sample included 10,691 youths (51% female): 8,550 Anglo Americans, 1,288 Mexican Americans, 409 Cuban Americans, and 444 Puerto Rican Americans. Actual ns varied somewhat across analyses, owing to missing data on other variables.

Measures

Racial-ethnic group. Respondents were classified into racial-ethnic groups on the basis of responses to the following four questions: (a) "What is your race? You may give more than one answer" (options: White, Black or African American, American Indian or Native American, Asian or Pacific Islander); (b) "Which one category best describes your racial background?" (options: White, Black or African American, American Indian or Native American, Asian or Pacific Islander); (c) "Are you of Hispanic or Latino origin?" (options: no, yes); and (d) "What is your Hispanic or Latino background?" (options: Mexican/Mexican American, Cuban/Cuban American, Puerto Rican, Central/South American, Other Hispanic). Respondents indicating Hispanic or Latino origin were included in the Latino group. Non-Latinos were classified as Anglo if they endorsed White as their only race or as the category that best described their racial background. Latinos were further divided according to their national origins. There were sufficiently large samples of Mexican, Cuban, and Puerto Rican American adolescents to permit subgroup analyses. Latinos who indicated another national origin or multiple national origins were excluded.

Depression. The Add Health study included a modified version of the 20-item CES-D. Slight differences in item wording were incorporated in the Add Health study. In most cases, the only difference was a shift from the first person to the second person (i.e., from "1" to "you"). In four cases, items were adapted to be more meaningful to adolescents (Items 7, II, 17 , and 20). Eighteen of the Add Health items appeared in one section of the survey. Adolescents were asked, "How often was each of these things true during the past week?" The response scale for those items ranged from 0 (*never or rarely*) to 3 (*most of the time or all of the time*). The other two items, "trouble falling asleep or staying asleep" and "frequent crying," were included in another section of the survey. Adolescents were asked, "Please tell me how often you have had each of the following conditions in the past 12 months." For these two items, the response scale was 0 (*never*) to 4 (*every day*). These items were recoded so that responses of 3 or 4 were assigned a value of 3.[1] Scale reliability was good: Alphas ranged from .84 to .87 across the four groups:

Self-esteem. A measure of self-esteem was included for the tests of .functional and scalar equivalence. There were six items, drawn from the Rosenberg Self-Esteem Inventory (Rosenberg, 1965) and similar scales (e.g., "You like yourself just the way you are."). Adolescents answered on a 5-point scale ranging from I (strongly agree) to 5 (strongly disagree). Items were reverse scored and averaged to form a total score, with higher scores indicating greater self-esteem. Alphas ranged from .83 to .87 across the four groups.

Demographic variables. Adolescents reported their gender, age, and generational status (i.e., whether they and their parents were born in the United States). Parents (in most cases, the mother) reported their educational attainment and whether any family members received public assistance (food stamps, housing subsidy, or Aid for Families With Dependent Children [AFDC]).

---

[1] The difference in response formats is of concern, especially because the specified time frame also varied. If the time frame influences responses substantially, then it might affect the size of factor loadings or result in these items forming their own factor. Fortunately, the factor loadings for the two items with an extended time frame did not differ markedly from those with the standard 1-week time frame.

Analytic Approach

The original four-factor structure identified by Radloff (1977) for Black and White adults was examined for Anglo American adolescents and for adolescents from the three Latino groups (Mexican, Cuban, and Puerto Rican). We first used CFA to examine the factor structure of the CES-D in each group; if the four-factor model showed acceptable fit in each group, configural invariance was supported. Where configural invariance between two groups was supported, we used multi-group CFAs to examine metric invariance (invariance of factor loadings across groups). Following recommendations outlined by Vandenberg and Lance (2000), we tested a series of nested models. In the first (unconstrained) model, the factor loadings and error variances were allowed to differ across groups; in the second model (metric invariance), factor loadings were constrained to be equal. A chi-square difference test was used to determine whether constraining the factor loadings to be equal resulted in a significant increase in chi-square (i.e., a significant decrement in model fit; Kline, 1998). If the difference in chi-square was nonsignificant, full metric invariance was supported. If the chi-square difference test was significant, additional models were tested to identify which factor loadings were invariant and which differed significantly.

To examine functional and scalar equivalence, we compared the associations between CES-D total scores and another theoretically relevant variable (self-esteem) across groups using multigroup structural equation modeling (SEM). Equivalent path coefficients (slopes) supported functional equivalence across groups; equivalent intercepts and slopes supported scalar equivalence (Knight & Hill, 1998).

In Add Health, schools were sampled and adolescents in those schools recruited for the study. Adolescents from the same school are more similar to each other than they are to adolescents from other schools (a clustering effect). In addition, certain groups were oversampled to ensure adequate sample sizes for analytic purposes. Failure to take these aspects of the sampling design into account leads to inaccurate point estimates and standard errors, biasing results toward finding differences between groups (Chantala & Tabor, 1999). Unless otherwise noted, we applied sample weights and accounted for the clustered design in all analyses to ensure that the results reflected the population figures. MPlus (Muthén & Muthén , 1998) was used for CFAs and SEMs because this program permits the use of both sampling weights and a cluster variable.[2]

To assess model fit, we used the Satorra-Bentler scaled (mean-adjusted) chi-square, a robust maximum-likelihood estimation technique; alpha was set at $p < .05$. The chi-square difference test was calculated following the method recommended by Satorra and Bentler (2001; see Muthén & Muthén , 1998). Obtaining a nonsignificant chi-square becomes increasingly unlikely with large sample sizes (Kline, 1998). Therefore, we focused on other indices of model fit that are less sensitive to sample size, including the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), and the standardized root-mean-square residual (SRMR). It is generally accepted that CFI values greater than .90 indicate adequate model fit (Kelloway, 1998; Kline, 1998; Maruyama, 1998), and we adopted this cutoff. For RMSEA and SRMR, the recommended criterion for good fit differs among sources, ranging from <.05 to <.10 (Hu & Bentler, 1995; Kelloway, 1998; Maruyama, 1998). We chose <.10 as the criterion for both indices; thus, only values less than .10 were deemed to show adequate fit.

Table 1

Demographic Characteristics of the Four Ethnic-Racial Groups

Characteristic	Anglo	Mexican	Cuban	Puerto Rican
n (unweighted)[a]	9,168	1,399	418	461
% women	51.03	49.15	51.79	50.33
% first generation	1.44	19.73	48.09	9.98
% college grade (parent)	24.95	6.81	12.91	11.88
% public assistance	8.29	17.28	25.56	31.83

Note. Values are based on unweighted data. Percentages are calculated with missing data excluded.

[a] Maximum ns. Actual ns may differ owing to missing data on individual variables.

Results

Demographic characteristics of the sample, based on unweighted data analyzed using SAS, are provided in Table 1. The ethnic subsamples did not differ in gender composition or mean age. However, parent education differed significantly across groups, χ^2 (3, $N = 10,029$) = 231.40, $p < .01$: Parents of Anglo adolescents were more likely to be college graduates than parents of youths in any Latino subgroup. The groups also differed significantly in public assistance, $\chi^2(3, N = 9,913) = 363.28$, $p < .01$: Parents of Anglo youths (8%) were least likely to report receiving public assistance, and parents of Puerto Rican youths were most likely to report receiving public assistance (32%). The four groups differed significantly in generational status, $\chi^2(6, N = 11,446) = 5,240.04$, $p < .01$. Anglo (1%) and Puerto Rican (10%) youths were far less likely to be first generation immigrants compared with Mexican (20%) and Cuban (48%) youths. The majority of Anglo youths (93%) and Puerto Rican youths (52%) were at least third generation, whereas the corresponding percentages for other Latino groups were lower (Mexican, 35%; Cuban, 3%).

Configural Invariance

Configural invariance was examined by testing the original four-factor solution (Radloff, 1977) in each of the four ethnic groups. As shown in Table 2, the four-factor solution fit well for Anglo American (CFI = .95, RMSEA = .04) and Mexican American adolescents (CFI = .92, RMSEA = .05), providing evidence of configural invariance across these two groups. However, the four-factor solution did not fit adequately for Cuban and Puerto Rican American youths alone, even after allowing errors to correlate[3] (CFIs = .76 and .83, RMSEAs = .18 and .08, respectively). .The poor model fit indicated a lack of configural invariance between these groups and the other two.

The four-factor solution for Anglo and Mexican American youths is depicted in Figure 1. The four factors included Negative Affect, Positive Affect (reverse scored), Interpersonal Aspects, and Somatic Symptoms. Factor loadings and factor intercorrelations were all significant for both groups. Factor intercorrelations were generally similar for the two groups, although the correlation between Negative Affect and Interpersonal Aspects was

---

[2] We used the complex sample modeling described in the Mplus manual (Type = complex). According to the manual, this approach "allows the use of sampling weights and a cluster variable. ..The standard errors and tests of model fit take into account the weights as well as the nonindependence of observations due to clustering" (Muthén & Muthén , 1998, p. 285).

[3] Mplus does not provide standard modification indices when incorporating design effects into the analyses. However, we examined derivatives and added correlated errors in the CFAs to improve model fit. We correlated errors only for items that were conceptually similar or were located next to each other on the survey.

somewhat higher among Mexican Americans than Anglo Americans. In both groups, correlations between the Negative Affect and Somatic Symptoms factors were large, whereas those between other pairs of factors were moderate in size.

Metric Invariance Tests for Anglo and Mexican Americans

A multigroup CF A was used to compare factor loadings for Anglo and Mexican Americans, the only two groups that showed configural invariance. A model in which factor loadings were constrained to be equal across the two groups was compared with one in which loadings were free to vary using a chi-square difference test calculated for use with the Satorra-Bentler adjusted chi-square (Muthén & Muthén , 1998). Factors were allowed to correlate freely in both models; these correlations are reported in Figure 1. The unconstrained two-group model showed good fit, χ^2 (325) = 3,076.17, $p < .01$, with a scaling correction factor of 1.084(CFI = .94, RMSEA = .04, SRMR = .03). The constrained model also showed good fit, χ^2 (341) = 3,139.44, $p < .01$, with a scaling correction factor of 1.097 (CFI = .94, RMSEA = .04, SRMR = .04). However, the difference in chi-square between the constrained and unconstrained models was significant, adjusted $\chi^2_{\text{difference}}$ (16) = 80.37, $p < .05$, so full metric invariance was not supported. We then tested for partial invariance by sequentially freeing the factor loadings that diverged most for the two groups until the chi-square difference test indicated that freeing additional loadings did not result in a significant decrease in chi-square (Vandenberg & Lance, 2000). As shown in Table 3, factor loadings for all but 3 of the 20 items could be constrained to be equal, providing strong evidence of partial invariance.[4] The fit of this model was good, χ^2 (338) = 3,070.37, $p < .01$, with a scaling correction factor of 1.093 (CFI = .94, RMSEA = .04, SRMR = .03); the adjusted chi-square difference test was not significant, χ^2(13) = 16.19, $p > .05$. The three CES-D items that differed for Anglo and Mexican Americans were "thought your life was a failure" and "felt fearful" from the Negative Affect factor and "enjoyed life" from the Positive Affect factor. In all three cases, the loadings were significantly higher for Mexican American youths than for Anglo youths.[5]

Table 2

Fit Indices for the Four-Factor Model of the Center for Epidemiologic Studies Depression Scale in the Four Groups

Group	χ^2	df	CFI	RMSEA	SRMR	n
Anglo	2,852.10*	164	.95	.04	.03	8,517
Mexican	609.14*	161	.92	.05	.04	1,279
Cuban	2,372.32*	164	.76	.18	.08	405
Puerto Rican	603.31 *	164	.83	.08	.06	444

Note. The four-factor model for Mexican Americans includes three pairs of correlated error terms. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

* $p < .001$.
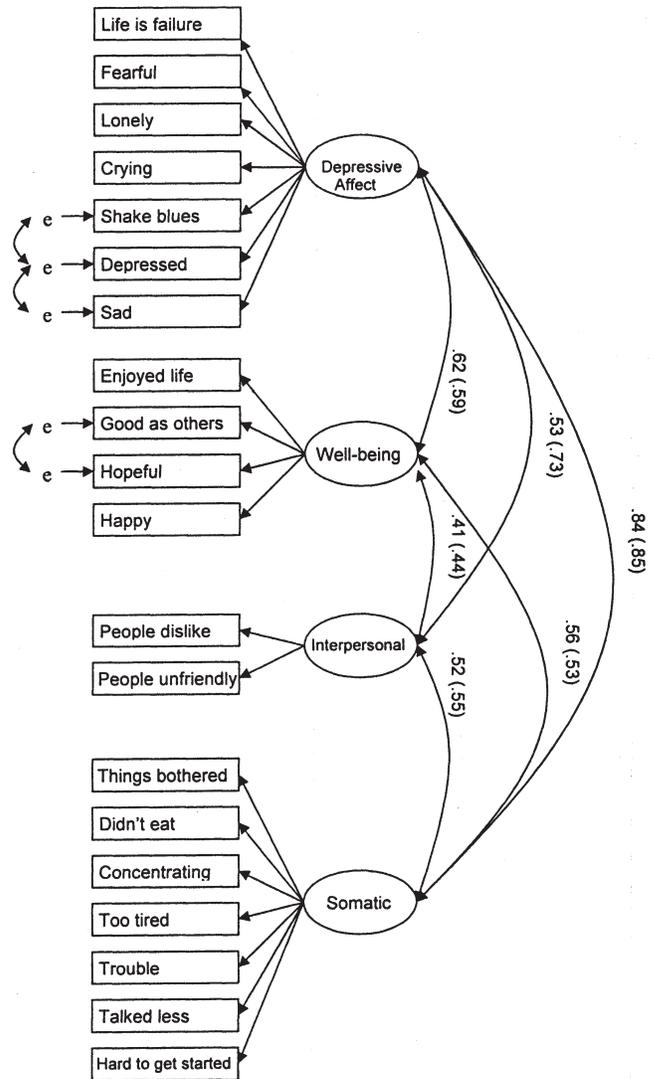
Figure 1. Four-factor model for Anglo and Mexican Americans. Factor correlations for Mexican Americans are in parentheses. Correlated error terms are for Mexican Americans only. e = error.

[4] The recommended approach for multi group CFAs is to standardize the latent factors by setting one indicator per factor to 1.0 (Kline, 1998). Thus, when testing for metric invariance, one indicator per factor is declared invariant by default. To test the invariance of these indicators, we ran a second set of analyses using a different item to standardize each latent factor. In all cases, the items that were originally used to standardize each factor were invariant in the second set of analyses. In addition, the results were substantially the same. In both sets of analyses, three items were noninvariant. Two of these items were the same in both cases ("failure" and "fearful"); however, the third noninvariant item differed; It was "enjoyed life" in the first analysis and "helpless" in the second.

[5] Although identical baseline models are not required for two-group CFAs (Bollen, 1989; Byrne, 1994), the CFA was also run with correlated error terms added for Anglo youths to match those required for Mexican youths. This change yielded a single noninvariant item ("failure") instead of three but did not alter the overall pattern of results.

Table 3

Unstandardized (and Standardized) Factor Loadings Showing Partial Metric Invariance for Anglo and Mexican American Adolescents on the Center for Epidemiologic Studies Depression Scale (CES-D)

CES-D factor and item	Anglo American ($n = 8{,}517$)	Mexican American ($n = 1{,}279$)
Depressive Affect		
Could not shake blues	0.84 (.75)	0.84 (.66)
Felt depressed	1.19 (.82)	1.19 (.74)
Thought life a failure	**0.52 (.54)**	**0.82 (.62)**
Felt fearful	**0.53 (.48)**	**0.67 (.55)**
Felt lonely	0.93 (.69)	0.93 (.65)
Frequent crying	0.66 (.50)	0.66 (.47)
Felt sad	1.00 (.75)	1.00 (.68)
Well-Being		
As good as others	0.87 (.54)	0.87 (.42)
Hopeful about future	1.01 (.54)	1.01 (.45)
Happy	1.15 (.74)	1.15 (.62)
Enjoyed life	**1.28 (.76)**	**1.55 (.76)**
Somatic Symptoms		
Bothered by things	1.02 (.59)	1.02 (.60)
Didn't feel like eating	0.90 (.51)	0.90 (.50)
Trouble concentrating	1.17 (.58)	1.17 (.59)
Too tired to do things	0.99 (.53)	0.99 (.54)
Trouble sleeping	1.02 (.39)	1.02 (.40)
Talked less than usual	0.69 (.41)	0.69 (.38)
Hard to get started	0.87 (.49)	0.87 (.51)
Interpersonal Issues		
Felt people disliked me	1.38 (.85)	1.38 (.76)
People were unfriendly	0.74 (.61)	0.74 (.58)

Note. The loading for the first item of each factor was initially set to 1.00 to standardize the metric. To obtain actual loadings for those items, we reran the analyses using another invariant item to standardize the metric. Noninvariant loadings appear in bold.

Follow-Up Analyses of Factor Structure for Cuban and Puerto Rican Americans

Generational status. Given the differing proportions of first, second, and third generation youths in the four ethnic groups, it was possible that differences in acculturation contributed to the lack of configural invariance found for Cuban and Puerto Rican Americans. The Add Health data set does not contain measures of acculturation, so we focused on the potential impact of generational status. In follow-up CFAs, we compared first and second generation Anglo Americans to first and second generation Cuban and Puerto Rican Americans, respectively (sample sizes in the Latino subgroups did not permit comparisons of first generation youths). Using this more restricted sample, the four-factor structure still fit for Anglo Americans but not for Puerto Rican and Cuban Americans. Thus, generational status does not appear to explain the lack of configural invariance.

Alternative factor structures. Research with Latino adults has sometimes shown a three-factor structure for the CES-D, and three factors also emerged among Hispanic youths who completed a 12-item version of the CES-D (Roberts, 1992). Therefore, we tested a three-factor confirmatory model for Cuban and Puerto Rican Americans, combining the depressed affect and the somaticizing symptoms into one factor. The fit indices suggested that this model did not fit for either Cuban Americans, χ^2 (167) = 2,635.79, $p < .01$ (CFI = .73, RMSEA = .19, SRMR = .08), or Puerto Rican Americans, χ^2 (167) = 644.01, $p < .01$ (CFI = .82, RMSEA = .08, SRMR = .06). Models for Puerto Rican and Cuban Americans that were based on results of the EFA reported by Guarnaccia et al. (1989) were also tested. However, the fit of these models was also poor among both Cuban Americans, χ^2 (163) = 2,736.67, $p < .01$ (CFI = .72, RMSEA = .20, SRMR = .09), and Puerto Rican Americans, χ^2 (165) = 762.53, $p < .01$ (CFI = .77, RMSEA = .09, SRMR = .07).

Finally, to provide descriptive information on the factor structure for Cuban and Puerto Rican American youths, we conducted an EFA of the CES-D items in each of these groups. The EFA was conducted in SAS (SAS Institute, 1999) with maximum-likelihood estimation; thus, the analyses included weights but the clustering effect was not taken into account.[6] Results are shown in Tables 4 and 5. A four-factor structure was found for Puerto Rican Americans (see Table 4). One factor appeared to reflect a multifaceted Negative Affect factor (11 items) that included items indexing sadness and anxiety plus two somatic items, a second corresponded to the typical Positive Affect factor (four items), a third was a Somatic factor characterized by low energy and poor concentration (three items), and the fourth was a modified Interpersonal factor (three items). One item ("loneliness") cross-loaded on two factors. For Cuban Americans, a five-factor solution emerged (see Table 5). The results were difficult to interpret because different kinds of symptoms loaded on the factors and because several items cross-loaded (e.g., "sadness" loaded on three factors). There appeared to be a Depressed Affect-Loneliness factor, a Somatic Complaints-Anxiety factor, and a Modified Well-Being factor (five items); the other factors were even less clear cut (e.g., interpersonal issues combined with "sad" and "bothered"; "bothered" combined with "hopeless" and "quiet"). These results suggest that Cuban (and, to some extent, Puerto Rican) youths do not show the distinct dimensions of depression found in other groups.

Functional and Scalar Equivalence

We examined functional and scalar equivalence using SEM. In a four-group SEM, self-esteem and control variables (child gender, age, parental education, and receipt of public assistance) were used to predict CES-D scale scores. Because group differences in slopes and intercepts can reflect nonequivalence in either the predictor or criterion variable, it was necessary to examine the measure equivalence of the self-esteem variable. CFAs (available from Lisa J. Crockett) supported configural invariance and full metric invariance between Anglo Americans and both Mexican and Puerto Rican Americans. (However, for Anglo and Cuban Americans, only configural invariance was found.) Thus, the self-esteem measure was a good candidate for examining the cross-ethnic functional and scalar equivalence of the CES-D between Anglos and two of the three Latino subgroups. It should be kept in mind that for Cubans and Anglos, lack of functional or scalar equivalence could result from nonequivalence of either the CES-D or the self-esteem scale.

[6] Mplus requires that the number of factors be specified and that one item on each factor be set to equal 1 in factor analyses that incorporate both weights and clustering effects; rather than introduce these constraints on the model, we ran the EFAs in SAS.

Table 4

Factor Loadings From an Exploratory Factor Analysis for Puerto Rican American Adolescents (n = 444)

CES-D item	Factor 1	Factor 2	Factor 3	Factor 4
Could not shake blues	**.85**	.07	−.17	.04
Felt depressed	**.63**	.15	.17	.05
Thought life a failure	**.39**	.19	.10	.04
Felt fearful	**.39**	−.02	.06	.12
Felt lonely	**.56**	−.05	−.05	.39
Frequent crying	**.39**	−.07	.11	.00
Felt sad	**.73**	−.05	.05	.08
As good as others	.03	**.43**	.27	.05
Hopeful about future	−.03	**.66**	−.10	.08
Happy	.03	**.69**	−.04	.06
Enjoyed life	−.09	**.74**	.03	.20
Bothered by things	**.47**	.10	.06	.00
Didn't feel like eating	**.33**	.24	.14	.16
Trouble concentrating	.28	.00	**.43**	.08
Too tired to do things	.00	−.01	**.72**	.03
Trouble sleeping	**.35**	−.13	.13	.07
Talked less than usual	**.37**	−.07	−.02	.21
Hard to get started	.23	−.03	**.40**	.12
Felt people disliked me	.12	.00	.14	**.64**
People were unfriendly	−.09	.06	−.07	**.68**

Note. Factor loadings of .30 or higher appear in bold. CES-D = Center for Epidemiologic Studies Depression Scale.

The four-group structural model is depicted in Figure 2, which shows the unconstrained path coefficients for each of the four groups. The exogenous variables were allowed to correlate, but, for simplicity, the correlations are not shown. The unconstrained model indicated predictable main effects for gender, age, parental education, and public assistance for most groups, although the effects were small and not always significant.

The model shown in Figure 2, in which no paths were constrained, was saturated and yielded a chi-square of 0.0 and a CFI of 1.0. Thus, for the chi-square difference tests, we estimated an alternative baseline model in which the path between one control variable (child age) and depression was constrained to be equal across groups. This model showed an acceptable fit, $\chi^2(3)$ = 65.56, $p < .01$, with a scaling correction factor of 0.49 (CFI = .98, RMSEA = .096, SRMR = .01). To examine functional equivalence, we compared this baseline model with a model in which the paths from child age and self-esteem to depression were each constrained to be equal across groups. The chi-square difference test was not significant, adjusted $\chi^2_{\text{difference}}(3) = 7.19$, $p > .05$, supporting functional equivalence. To examine scalar equivalence, we compared this model with a model in which the slope and intercept for self-esteem (as well as the slope for child age) were constrained to be equal across groups. The chi-square difference test, calculated for the Satorra-Bentler adjusted chi-square, was not significant even in the latter case, adjusted $\chi^2_{\text{difference}}(3) = 6.70$, $p > .05$, supporting scalar equivalence of the CES-D.[7]

Discussion

The goal of the present study was to investigate the cross-ethnic equivalence of the CES-D across representative samples of Anglo and Latino adolescents, including Mexican, Cuban, and Puerto Rican youths. To this end, we examined configural and metric invariance at the item level, as well as functional and scalar equivalence of CES-D scale scores, Results provided mixed support for measurement equivalence of the CES-D for adolescents, The patterns varied across the types of equivalence tested and differed for different pairs of ethnic groups.

Results of CFAs supported the original four-factor structure of the CES-D (Radloff, 1977) among Anglo and Mexican American adolescents but not among Cuban and Puerto Rican youths. The results for Mexican and Anglo Americans indicate that symptoms included in the CES-D cluster in the same way in both groups, reflecting the same underlying dimensions. Thus, it appears likely that CES-D scores represent the same construct in these groups (Ghorpade et al., 1999). Even so, full metric invariance was not supported, as up to three items loaded more strongly on the Negative Affect factor for Mexican Americans compared with Anglo youths. Thus, although the two groups responded in similar ways to depressive symptoms on the CES-D, some items appeared to be more salient indicators of depression for Mexican youths than for Anglo youths.

A lack of full metric invariance could increase the risk of classification errors when the CES-D is used for screening in epidemiological studies, resulting in misestimation of the prevalence of depression symptoms among Mexican American adolescents (Posner et al., 2001). However, with only 3 noninvariant items out of 20, the amount of misclassification may be small. We attempted to estimate the degree of classification error by calculating who in the present sample would meet standard cutoffs for depression using the ful120-item CES-D and a 17-item version in which the three noninvariant items were excluded; the cutoff score was adjusted accordingly. On the basis of weighted analyses conducted in SAS, 25.2% of Anglo and 35.6% of Mexican Americans met the cutoff with the 20-item version, whereas the comparable percentages with the 17 -item version were 26.4% and 37.4%. Thus, compared with the 17-item version, the 20-item version would underestimate depression in 10/0-2% of Anglo and Mexican American youths. If this is an acceptable error rate given one's research goals, then it should be possible to use the CES-D in studies of Mexican American adolescents.

The findings for Anglo and Mexican Americans are consistent with previous EFAs of the CES-D, in which factor structures for Anglo and Mexican Americans were generally similar but minor differences emerged in factor loadings or factor structure. For example, Garcia and Marks (1989) found that "lonely," "sad," and "crying" loaded together on the Negative Affect factor for Mexican American adults, unlike the pattern typically reported for Anglo samples. Similarly, Roberts (1980) found that some negative affect and somatic items loaded together for Mexican American adults but not for African Americans and Anglo Americans.

---

[7] Functional and scalar equivalence for Anglo and Mexican Americans were also examined in two-group SEMs in which self-esteem and the CES-D were modeled as latent variables, with item scores serving as observed indicators. For the CES-D, factor loadings were constrained to be equal across groups, except for the three noninvariant items; three pairs of correlated errors were included for the Mexican Americans. For self-esteem, all factor loadings were constrained to be equal (with one correlated error term). The pattern of results was unchanged; on the basis of chi-square difference tests. all four groups continued to show functional and scalar equivalence.
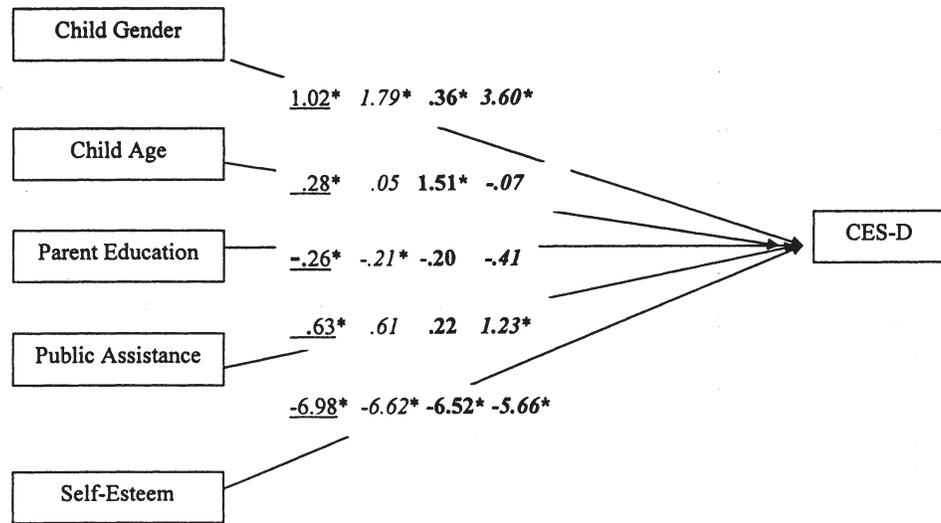
Figure 2. Structural equation model predicting Center for Epidemiologic Studies Depression Scale (CES-D) total scores from adolescent self-esteem and controls. For each path, coefficients are provided for Anglo, Mexican, Cuban, and Puerto Rican Americans, respectively. Underlined numbers = Anglo Americans (*n* = 7,426); italicized numbers = Mexican Americans (*n* = 975); bold numbers = Cuban Americans (*n* = 345); italicized bold numbers = Puerto Rican Americans; asterisks indicate significant paths within ethnic groups.

Using a 12-item version of the CES-D, Roberts (1992) found that for Mexican American youths and other Hispanics, but not other groups, a few somatic items tended to cross-load on the Somatic and Negative Affect factors. Combined with the results of the present study, it appears that the differences between Anglo and Mexican adolescents in their depression symptomatology are very subtle.

The findings for Anglo youths are also noteworthy because they show that the factor structure identified in White and Black adults (e.g., Radloff, 1977) also applies to contemporary White adolescents. Our results, based on CFAs, support those of Roberts et al. (1990), who used EFA to replicate the original four-factor solution in a sample of non-Hispanic White youths. Although we could not test for metric invariance between adolescents and

Table 5
Factor Loadings From an Exploratory Factor Analysis for Cuban American Adolescents (n = 405)

CES-D item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Could not shake blues	**.66**	.14	.08	-.10	.08
Felt depressed	**.79**	.01	.22	-.11	.07
Thought life a failure	**.48**	.23	-.15	.15	.09
Felt fearful	.19	**.39**	.23	-.04	-.11
Felt lonely	.50	-.18	**.46**	.22	-.08
Frequent crying	.91	-.00	-.29	-.07	.05
Felt sad	.35	-.05	**.51**	.03	.34
As good as others	.10	-.10	.06	.06	**.62**
Hopeful about future	-.24	.02	-.12	**.70**	.42
Happy	.21	.26	.01	.15	**.43**
Enjoyed life	.09	.27	.19	.13	**.34**
Bothered by things	.18	.18	.37	**.40**	-.13
Didn't feel like eating	.01	**.47**	.02	.02	.10
Trouble concentrating	**.37**	.29	.14	.11	-.16
Too tired to do things	.04	.76	.09	-.05	-.13
Trouble sleeping	.06	.10	-.23	**.30**	.02
Talked less than usual	.03	-.07	.09	**.52**	.02
Hard to get started	.03	**.51**	.18	.21	-.14
Felt people disliked me	-.04	.29	**.75**	-.22	.07
People were unfriendly	-.23	.18	**.80**	-.13	.06

Note. Factor loadings of .30 or higher appear in bold. CES-D = Center for Epidemiologic Studies Depression Scale.

adults because only adolescents were sampled, the fact that the four-factor structure identified for Black and White adults also fit the data for Anglo adolescents is consistent with configural invariance across generations.

In contrast, the present study did not support the established four-factor structure for either Puerto Rican or Cuban youths. These findings partially mirror previous research with adults, in which EFAs indicated different factor structures for Cuban, Puerto Rican, and Mexican Americans (Guarnaccia et al., 1989). Puerto Rican Americans showed a four-factor structure in which negative affect and somatic symptoms loaded on one factor, suggesting a co-occurrence of these symptoms and a blurring of the distinction between affective and somatic symptoms, as has been reported in some prior research with Latinos (Roberts, 1980, 1992). The Positive Affect and Interpersonal factors were largely intact, with some minor variations. For Cuban Americans, however, five factors emerged, none of which was readily interpretable. Research with adults has also found that Cuban Americans show a unique factor structure for the CES-D. For example, Guarnaccia et al. (1989) reported a factor for Cuban Americans that appeared to reflect feelings of rejection by and isolation from the larger society. The different factor structures observed for Cuban and Puerto Rican Americans suggest that the symptom clusters experienced by these groups may differ from those experienced by Anglos and Mexicans.

The observed differences in factor structure could stem from multiple sources. Lack of configural invariance is often interpreted as evidence that two groups do not share a common cognitive frame of reference for the construct being measured (Ghorpade et al., 1999; Vandenberg & Lance, 2000). If so, one possible explanation for the present findings is that Cuban and Puerto Rican Americans have somewhat different concepts of depression than Anglo Americans, leading them to experience (or at least report) different patterns of symptoms. Ethnic differences in concepts would be consistent with the notion that concepts of mental health and illness are culturally derived (Choi, 2002; James & Prilleltensky, 2002). For example, James and Prilleltensky (2002) suggested that "social values shape the conceptualization and the social construction of mental health" (p. 1137) and that "norms affect the group's conceptualization and experience of mental health" (p. 1144).

Alternatively, the distinct factor structures of Cuban and Puerto Rican Americans could be a response to aspects of their social context, including cultural, economic, and other environmental factors. 'According to Kleinman's (1986) sociosomatic formulation, a person's context (cultural, social, economic) "privileges" different symptom clusters, making them more prevalent and more central indicators of distress or disorder in that context (James & Prilleltensky, 2002). Such differences need not involve different concepts of disorders but simply different values, norms, or experiences. Regardless of whether the distinct factor structures identified in the present study reflect different concepts of depression among Puerto Rican and Cuban Americans or just differential symptom expression, CES-D scores may not be comparable across groups.

The divergent results for Cuban and Puerto Rican Americans do not appear to be attributable to differences in generational status. However, sample sizes of Cuban and Puerto Rican Americans did not permit a precise test of this possibility. Other research using Add Health data has identified differences in CES-D factor structure that appear to be associated with generational status among Latinos (Perreira, Deeb-Sossa, Harris, & Bollen, 2003). That study did not examine distinct subgroups of Latinos, so the role of generational status for Cuban and Puerto Rican youths (and the role of acculturation more broadly) remains an important direction for future research.

The lack of invariance in the present study could also reflect differences in sample size: Smaller numbers of Cuban and Puerto Rican youths could have resulted in less stable estimates. However, the sample size for each group was greater than 400, which should have been sufficient for obtaining reliable estimates. It is also possible that Cuban and Puerto Rican youths are more heterogeneous than Mexican or Anglo youths and hence more difficult to capture with a single factor structure. In the present case, the parents of Puerto Rican and Cuban youths reported more education than parents of Mexican Americans but also more public assistance, suggesting greater heterogeneity with respect to socioeconomic status.[8] In light of these differences, the present results should be replicated in other studies that can take such variability into account. Still, the lack of configural invariance for these youths suggests that they express different symptom clusters than either Anglo or Mexican American youths. Additional research is needed to determine the range of depression symptoms they experience, the dimensions underlying those clusters, and what factors account for group differences in symptom clusters. Qualitative studies may be needed to elucidate the meaning of specific symptoms (CES-D items) for Cuban and Puerto Rican adolescents and to determine whether these youths hold distinct concepts of depression.

In contrast to the findings regarding configural and metric invariance, the SEMs provided evidence of functional and scalar equivalence across the four ethnic groups. Similar relations between CES-D total scores and self-esteem were found in all four groups, with socioeconomic indicators controlled. Thus, the construct measured by the CES-D appeared to function similarly for Anglo, Mexican, Cuban, and Puerto Rican youths.

Taken together, the results suggest that CES-D scores capture a form of distress that functions similarly in the four ethnic groups. Thus, it might be possible to use the CES-D as a general measure of negative affect in studies of the correlates of distress. At the same time, individual CES-D items did not operate in the same way, especially for Cuban and Puerto Rican youths, so cross-ethnic comparisons involving Latino youths could potentially yield inaccurate results. The potential problems of using the CES-D with Puerto Rican and Cuban youths demonstrated here suggest that we currently run the risk of both false positives and false negatives in screening for depression in some Latino youths. The risk of misclassification is greatest with Cuban and Puerto Rican youths, who may not share the same depression symptomatology as Anglo and Mexican American youths. Given the distinct results for the different Latino subgroups, studies that pool Latinos of different national origins appear ill advised, as do studies of depression using mixed samples of Latino and Anglo youths. This underscores the importance of understanding the meaning and expression of depression symptomatology among Latino subgroups in the United States as well as the need to develop ways to measure depression accurately in these groups.

[8] Higher levels of public assistance could also be due to greater likelihood of being eligible because Puerto Rican Americans are citizens and Cuban Americans usually start with refugee status.

The nonequivalent factor structures found for Cuban and Puerto Rican adolescents have implications for screening, assessment, and treatment. First, when the CES-D is used to screen for depression, researchers should recognize the risk that people from particular ethic groups or subgroups are more likely to be misclassified in epidemiological studies, leading to misestimations of prevalence rates. Such inaccuracies could affect public policy decisions on the local or national level, resulting in a faulty distribution of resources. To avoid this, screening instruments that are equivalent across ethnic groups are needed. Second, there are important clinical assessment and treatment implications of the finding that different symptoms appear to be more salient for some groups than others. Because some groups appear to experience different symptom clusters, clinicians who work with those groups may need to adjust their own concepts of depression to permit appropriate diagnosis and treatment. In other words, we may need to view depression as a "fuzzy concept" or a family of overlapping concepts rather than as a single disorder that presents in a uniform way. At the very least, it is important that practitioners know that depression may present differently across different ethnic groups (Minsky et al., 2003; Sue, 1998). According to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, "culture I can influence the experience and communication of symptoms of I depression. Underdiagnosis or misdiagnosis can be reduced by ! being alert to ethnic and cultural specificity in the presenting complaints of a Major Depressive Episode" (American Psychiatric .Association, 1994, p. 324).

Although the present study used a national data set with representative samples of Latino and Anglo youths, certain limitations apply. The version of the CES-D included in the Add Health study contained several modifications that could have influenced the results. However, the replication of the original factor structure among Anglo youths bolsters confidence in the version used. In addition, the Add Health sample was based on in-school youths and may not adequately represent the adolescents who are most at risk for depression. Furthermore, the analytic sample we used was restricted because of our desire to compare groups of youths with clear ethnic and racial affiliations. This led us to exclude youths with multiple affiliations (e.g., biracial youths who did not identify themselves primarily as White). Finally, as noted earlier, sample sizes differed across the ethnic groups, and, in some Latino subgroups, sample sizes did not permit a full examination of generational status effects or subgroup differences associated with other sociodemographic factors.

Nonetheless, the present results add to the growing recognition that established measures may not be equivalent across various racial-ethnic groups and that it may be misleading to apply instruments developed on one population to other populations without clear evidence of measurement equivalence. Studies of measurement equivalence of the CES-D should be extended to

other ethnic groups and subgroups. Where feasible, it would also be useful to examine within subgroup differences related to socioeconomic status or other important status characteristics.[9] Additional measures should be included to examine functional and scalar equivalence of the CES-D. Finally, the study of measurement equivalence should be extended to other mental health measures to enable accurate comparisons among diverse ethnic and racial groups.

References

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Aunola, K., Stat tin, H., & Nurmi, J. E. (2000). Adolescents' achievement strategies, school adjustment, and externalizing and internalizing problem behaviors. Journal of Youth and Adolescence, 29, 389–406.

Bearman, P. S., Jones, J., & Udry, J. R. (1997). The National Longitudinal Study of Adolescent Health: Research design. Retrieved January 30, 2001' from http://www.cpc.unc.edu/addhealth

Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., Kaufman, J., Dahl, R. E., et al. (1996). Childhood and adolescent depression: A review of the past 10 years. Pt. 1. Journal of the American Academy of Child and Adolescent Psychiatry, 35, 1427–1439.

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming. Thousand Oaks, CA: Sage.

Chantala, K., & Tabor, J. (1999). Strategies to perform a design-based analysis using the Add Health data. Retrieved July 27, 2001, from http://www.cpc.unc.edu/addhealth

Choi, H. (2002). Understanding adolescent depression in ethnocultural context. Advances in Nursing Science, 25, 71–86.

Fountoulakis, K., Iacovides, A., Kleanthous, S., Samolis, S., Kaprinis, S. G., Sitzoglou, K., et al. (2001). Reliability, validity, and psychometric properties of the Greek translation of the Center for Epidemiological Studies-Depression (CES-D) Scale. BMC Psychiatry, 1, 1–20. Retrieved June 18, 2002, from http://www.biomedcentral.com/1471-224x/.l/3

Garcia, M., & Marks, G. (1989). Depressive symptomatology among Mexican American adults: An examination with the CES-D scale. Psychiatry Research, 27, 137–148.

Garrison, C. C., Waller, J. L., Cuffe, S. P., & McKeown, R. E. (1997). Incidence of major depressive disorder and dysthymia in young adolescents. Journal of the American Academy of Child and Adolescent Psychiatry, 36, 458–465.

Ghorpade, J., Hattrup, K., & Lackritz, J. R. (1999). The use of personality measures in cross-cultural research: A test of three personality scales across two countries. Journal of Applied Psychology, 84, 670–679.

Golding, J. M., & Aneshensel, C. S. (1989). Factor structure of the CES-D scale among Mexican Americans and non-Hispanic Whites: Psychological assessment. Journal of Consulting and Clinical Psychology, I, 163–168.

Guarnaccia, P. J., Angel, R., & Worobey, J. L. (1989). The factor structure of the CES-D in the Hispanic Health and Nutrition Examination Survey: The influences of ethnicity, gender, and language. Social Science and Medicine, 29, 85–94.

Hu, L.T., & Bentler, P. M. (1995). Evaluating model fit. In R.H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 76–99). Thousand Oaks, CA: Sage.

Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology. Journal of Cross-Cultural Psychology, 16, 131–152.

---

[9] In many cases, it may not be practical to investigate the equivalence-invariance of all study measures across all ethnic, racial, and socioeconomic groups. The importance of doing so depends on the nature of the research question and the way in which results will be used. It is most important when the negative consequences of mismeasurement are greatest-when inaccuracies could lead to misguided public policies or unnecessary burdens to individuals. Mental health is an arena in which the stakes appear to be high, increasing the urgency of examining measurement equivalence.

James, S., & Prilleltensky, I. (2002). Cultural diversity and mental health: Towards integrative practice. Clinical Psychology Review, 22, 1133–1154.

Kelloway, E. K. (1998). Using LISREL for structural equation modeling. Thousand Oaks, CA: Sage.

Kleinman, A. (1986). Culture and depression. Berkeley: University of California Press.

Kline, R. B. (1998). Principles and practice of structural equation modeling. New York: Guilford Press.

Knight, G. P., & Hill, N. E. (1998). Measurement equivalence in research involving minority adolescents. In V. McLoyd & L. Steinberg (Eds.), Studying minority adolescents: Conceptual, methodological and theoretical issues (pp. 183–210). Hillsdale, NJ: Erlbaum.

Knight, G. P., Virdin, L. M., Ocampo, K. A., & Roosa, M. (1994). An examination of the cross-ethnic equivalence of measures of negative life events and mental health among Hispanic and Anglo American children. American Journal of Community Psychology, 22, 767–783.

Knight, G. P., Virdin, L. M., & Roosa, M. (1992). Socialization and family correlates of mental health outcomes among Hispanic and Anglo American children: Consideration of cross-ethnic scalar equivalence. Child Development, 65, 212–224.

Liang, J., Tran, T. V., Krause, N., & Markides, K. S. (1989). Generational differences in the structure of the CES-D Scale in Mexican Americans. Journal of Gerontology, 44, 110–120.

Maruyama, G. M. (1998). Basics of structural equation modeling. Thousand Oaks, CA: Sage.

McArdle, J. J., Johnson, R. C., Hishinuma, E. S., Miyamoto, R. H., & Andrade, N. N. (2001). Structural equation modeling of group differences in CES-D ratings of native Hawaiian and non-Hawaiian high school students. Journal of Adolescent Research, 16, 108–149.

Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. Archives of General Psychiatry, 60, 637–644.

Muthén, L. K., & Muthén, B. 0. (1998). Mplus user's guide. Los Angeles: Author.

Narrow, W. E., Rae, D. S., Moscicki, E. K., Locke, B. Z., & Regier, D. A. (1990). Depression among Cuban Americans: The Hispanic Health and Nutrition Examination Survey. Social Psychiatry and Psychiatric Epidemiology, 25, 260–268.

Ohannessian, C. M., Lerner, R. M., Lerner, J. V., & von Eye, A. (1999). Does self-competence predict gender differences in adolescent depression and anxiety? Journal of Adolescence, 22, 397–411.

Orme, J. G., Reis, J., & Hen, E. J. (1986). Factorial and discriminant validity of the Center for Epidemiological Studies Depression (CES-D) Scale. Journal of Clinical Psychology, 42, 28–33.

Perreira, K. M., Deeb-Sossa, N., Harris, K: M., & Bollen, K. (2003, July). What are we measuring? An evaluation of the CES-D across race, ethnicity, and immigrant generation. Paper presented at the Add Health Users Workshop, Bethesda, MD.

Posner, S. F., Stewart, A. L., Marin, G., & Perez-Stable, E. J. (2001). Factor variability of the Center for Epidemiological Studies Depression Scale (CES-D) among urban Latinos. Ethnicity and Health, 6, 137–144.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, I, 385–401.

Roberts, R. E. (1980). Prevalence of psychological distress among Mexican Americans. Journal of Health and Social Behavior, 21, 130–145.

Roberts, R. E. (1992). Manifestation of depressive symptoms among adolescents: A comparison of Mexican Americans with the majority and other minority populations. Journal of Nervous and Mental Disease, 180, 627–633.

Roberts, R. E., Andrews, J. A., Lewinsohn, P. M., & Hops, H. (1990). Assessment of depression in adolescents using the Center for Epidemiologic Studies Depression Scale. Psychological Assessment, 2, 122–128.

Roberts, R. E., & Chen, Y. W. (1995). Depressive symptoms and suicidal ideation among Mexican-origin and Anglo adolescents. Journal of the American Academy of Child and Adolescent Psychiatry, 34, 81–90.

Roberts, R. E., Roberts, C. R., & Chen, Y. R. (1997). Ethnocultural differences in prevalence of adolescent depression. American Journal of Community Psychology, 25, 95–110.

Roberts, R. E., & Sobhan, M. (1992). Symptoms of depression in adolescence: A comparison of Anglo, African, and Hispanic Americans. Journal of Youth and Adolescence, 21,639–651.

Roberts, R. E., Vernon, S. W., & Rhoades, H. M. (1989). Effects of language and ethnic status on reliability and validity of the Center for Epidemiologic Studies-Depression Scale with psychiatric patients. Journal of Nervous and Mental Disease. 177. 581–592.

Rosenberg, M. (1965). Society and adolescent self-image. Princeton, NJ: Princeton University Press.

Rushton, J. L., Forcier, M., & Schectman, R. M. (2002). Epidemiology of depressive symptoms in the National Longitudinal Study of Adolescent Health. Journal of the American Academy of Child and Adolescent Psychiatry.41, 199–205.

Salgado de Snyder, A. N., Diaz-Perez, M. J., & Ojeda, V. D. (2000). The prevalence of nervios and associated symptomatology among inhabitants of Mexican rural communities. Culture, Medicine and Psychiatry, 24. 453–470.

SAS Institute. (1999). The SAS System: Version 8. Cary, NC: Author.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66, 507–514.

Seigel, J. M., Aneshensel, C. S., Taub, B., Cantwell, D. P., & Driscoll, A. K. (1998). Adolescent depressed mood in a multi-ethnic sample. Journal of Youth and Adolescence, 27,413–427.

Shaffer, D., Fisher, P., Dulkan, M. K., Davies, M., Piacentini, J., Schwab-Stone, M. E., et al. (1996). The NIMH Diagnostic Interview Schedule for Child, Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA study. Journal of the American Academy of Child and Adolescent Psychiatry. 35, 865–877.

Sue, S. (1998). In search of cultural competence in psychotherapy and counseling. American Psychologist. 53. 440–448.

Udry, J. R. (1998). The National Longitudinal Study of Adolescent Health (Add Health). Waves I and 11, 1994–1996. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

U.S. Department of Health and Human Services. (2001). Mental health: Culture. race. ethnicity. A supplement to mental health: A report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organization research. Organizational Research Methods, 3. 4–70.

van Dulman, M. H. M., Grotevant, H. D., Dunbar, N., Miller, B. C., Bayley, B., Christensen, M., & Fan, X. (2002). Connecting national survey data with DSM-IV criteria. Journal of Adolescent Health, 36, 475–481.

Vega, W. A., Khoury, E. L., Zimmerman, R. S., Gil, A. G., & Warheit, G. J. (1995). Cultural conflicts and problem behaviors of Latino adolescents in home and school environments. Journal of Community Psychology,23, 167–179.

Vega, W. A., & Rumbaut, R. G. (1991). Ethnic minorities and mental health. Annual Review of Sociology, 17, 351–383.

Weinberg, W. A., & Emslie, G. J. (1987). Depression and suicide in adolescents. International Pediatrics, 2, 54–159.